

WHAT IS CLAIMED IS:

1. A file system including at least one node having
a file server which processes files distributed and managed
5 in a plurality of physical disk devices, wherein a file ID
is defined for each of said files, said node comprising:

a file management table including records each
composed of a file ID and a logical disk ID of a logical
disk, said logical disk storing a file corresponding to
10 said file ID;

a logical disk management table including records
each composed of said logical disk ID and one or more I/O
paths, said one or more I/O paths being used for accessing
one or more physical disk devices corresponding to said
15 logical disk;

wherein upon receiving a request for accessing a
file specifying a file ID from a user, the file server
refers to said file management table; determines a logical
disk ID of a logical disk storing said file based on said
20 file ID; refers to said logical disk management table to
determine an I/O path for accessing a physical disk device
corresponding to said logical disk based on said logical
disk ID; and accesses the physical disk device by use of
the determined I/O path.

25 2. A file system including a plurality of nodes

which are connected to a network and which each has a file server therein, and a physical disk device commonly connected to at least two nodes included in the plurality of nodes, wherein said file system processes files which
5 are distributed and managed in said plurality of physical disk devices and a file ID is defined for each of said files, said plurality of nodes each comprising:

a file management table including records each composed of a file ID and a logical disk ID of a logical
10 disk, said logical disk storing a file corresponding to said file ID;

a logical disk management table including records each composed of said logical disk ID and one or more I/O paths, said one or more I/O paths being used for accessing
15 one or more physical disk devices corresponding to said logical disk;

wherein upon receiving a request for accessing a file specifying a file ID from a user, a file server refers to the file management table; determines a logical disk ID
20 of a logical disk storing said file based on said file ID; refers to said logical disk management table to determine an I/O path for accessing a physical disk device corresponding to said logical disk based on said logical disk ID; and accesses the physical disk device by use of
25 the determined I/O path.

3. The file system as claimed in claim 2, wherein information specifying said I/O path comprises a node number, an I/O interface number, and a disk controller number.

5 4. The file system as claimed in claim 3, wherein when the physical disk device corresponding to the logical disk ID determined based on said file ID is connected to a remote node, a file server in a local node transmits an access request to said remote node, and a file server in
10 said remote node which has received said access request accesses a given file stored in said physical disk device.

5. A file system including a plurality of nodes which are connected to a network and which each has a file server therein, and a physical disk device commonly
15 connected to at least two nodes included in the plurality of nodes, wherein said file system processes files which are distributed and managed in said plurality of physical disk devices and a file ID is defined for each of said files, said plurality of nodes comprising:

20 a first node having at least one disk device connected thereto which stores a mount construction file one of whose entries includes information for associating a mount point with one or more I/O paths for accessing a physical disk device,

25 wherein at system start-up, the file server in said

first node automatically sets a logical disk ID for the one or more I/O paths included in the entry in said mount construction file; registers a matching relationship between said logical disk ID and said one or more I/O paths in a logical disk management table; copies contents of said logical disk management table to logical disk management tables in all other nodes; and mounts a logical disk corresponding to said logical disk ID onto the mount point associated with said one or more I/O paths in said mount construction file; and

a plurality of second nodes each including:

a file management table including records each composed of a file ID and a logical disk ID of a logical disk, said logical disk storing a file corresponding to said file ID;

a logical disk management table including records each composed of said logical disk ID and one or more I/O paths, said one or more I/O paths being used for accessing one or more physical disk devices corresponding to said logical disk;

wherein upon receiving a request for accessing a file specifying a file ID from a user, a file server refers to said file management table; determines a logical disk ID of a logical disk storing said file based on said file ID; refers to said logical disk management table to determine

an I/O path for accessing a physical disk device corresponding to said logical disk based on said logical disk ID; and accesses the physical disk device by use of the determined I/O path.

5 6. The file system as claimed in claim 5, wherein said mount construction file includes availability information which is set for each I/O path and indicates whether the I/O path is available, and said logical disk management table includes status flags each indicating an
10 operational state of each I/O path registered therein, whereby at system start-up, the file server in said first node which performs mount operation sets "operational" for a status flag in the logical disk management table corresponding to an I/O path whose
15 availability information is set to "available" in said mount construction file and which is selected from the plurality of I/O paths included in the entry in said mount construction file; sets "standby" for each of status flags in said logical disk management table corresponding to
20 remaining I/O paths whose availability information is set to "available" in said mount construction file; and sets "unavailable" for each of status flags in said logical disk management table corresponding to I/O paths whose availability information is set to "unavailable" in said
25 mount construction file; whereas a file server in each of

said plurality of second nodes accesses a physical disk device using an operational I/O path whose status flag is set to "operational" in said logical disk management table in normal operation.

5 7. The file system as claimed in claim 6, wherein when the operational I/O path has become unavailable, a file server in a node which has detected the fault updates a logical disk management table in said node by setting "unavailable" for a status flag of said unavailable I/O
10 path, and setting "operational" for a status flag of an I/O path selected from among I/O paths whose status flag is set to "standby" and which are associated with a same logical disk ID with which said unavailable I/O path is associated so as to designate the selected I/O path as a new
15 operational I/O path, and then copies contents of said logical disk management table to logical disk management tables in all other nodes to switch from said unavailable I/O path to said new operational I/O path for accessing said physical disk device.

20 8. The file system as claimed in claim 7, wherein during a process of switching said I/O paths, a file server in a node included in the unavailable I/O path reserves an access request directed to the unavailable I/O path, and transfers the reserved access request to a node included in
25 the new operational I/O path after the process of switching

the I/O paths has been completed.

9. The file system as claimed in claim 7, wherein during a process of switching said I/O paths, when an access request issued from a file server to a node included in the unavailable I/O path has timed out, the file server refers to said logical disk management table to obtain an I/O path based on the logical disk ID again, and uses the obtained I/O path to access the physical disk device.

10. The file system as claimed in claim 7, wherein said plurality of nodes each have a buffer cache for temporarily holding data to be transferred between the node and the physical disk device, and during a process of switching the I/O paths, a file server in a node included in the unavailable I/O path communicates with a file server in a node included in the new operational I/O path so as to transfer contents of a buffer cache and said file management table which exist in a main memory of the node included in said unavailable I/O path, to the node included in said new operational I/O path, said contents of said buffer cache and said file management table being necessary to write back to the physical disk device.

11. The file system as claimed in claim 7, wherein a disk controller in said physical disk device comprises:

a disk cache for temporarily holding data to be transferred to and from a disk area; and

means for writing back data stored in a disk cache provided in another disk controller in said physical disk device to the disk area;

wherein during a process of switching the I/O paths,
5 a disk controller included in the new operational I/O path and provided in the physical disk device writes back data selected from data stored in a disk cache in another disk controller included in said unavailable I/O path and provided in said physical disk device, to said physical
10 disk device, which was being accessed using said currently unavailable I/O path, said selected data being necessary to write back to said physical disk device.

12. The file system as claimed in claim 7, wherein at the time of completing a process of switching the I/O
15 paths, the file server in said first node to which the disk device storing said mount construction file is connected updates said mount construction file, and rewrites availability information on said unavailable I/O path by information indicating "unavailable".

20 13. The file system as claimed in claim 7, wherein when the currently unavailable I/O path has become available again, a file server in one of said plurality of nodes updates a status flag of said I/O path registered in a logical disk management table in the node by changing the
25 status flag from "unavailable" to "standby" and copies said

updated contents of the logical disk management table to logical disk management tables in all other nodes, and the file server in the node to which the disk device storing the mount construction file is connected subsequently
5 rewrites availability information on said I/O path registered in said mount construction file by information indicating "available" so as to restore said I/O path in the system as a standby I/O path.

14. The file system as claimed in claim 6, wherein
10 when a fault has occurred in a node to which a physical disk device is connected, a file server in another node which has detected said node fault searches a logical disk management table in the another node to obtain the faulty I/O path determined based on a node number of the faulty
15 node and an I/O path selected from among I/O paths whose status flag is set to "standby" and which correspond to a logical disk ID included in said faulty I/O path and designate the selected I/O path as a new operational I/O path, and requests a file server in a node included in said
20 new operational I/O path to switch the two I/O paths, and said file server which has received said request updates a logical disk management table in its own node by setting a status flag of said faulty I/O path to "unavailable" and a status flag of said new operational I/O path to
25 "operational", then copies contents of said logical disk

management table to logical disk management tables in all other nodes, and switches from said faulty I/O path to said new operational I/O path for accessing said physical disk device.

5 15. The file system as claimed in claim 14, wherein during a process of switching the I/O paths, when an access request issued from a file server to a node included in said faulty I/O path has timed out, the file server refers to a logical disk management table to obtain an I/O path
10 based on the logical disk ID again, and uses the obtained I/O path to access the physical disk device again.

 16. The file system as claimed in claim 14, wherein the node to which said physical disk device is connected has a means for, regardless of conditions of the node,
15 reading out data in a memory provided in the node and transferring the read-out data to another node, and whereby during a process of switching the I/O paths, the node transfers contents of a buffer cache and said file management table which exist in a main memory of a node
20 included in the said faulty I/O path to a node included in said new operational I/O path by use of said means, said contents of said buffer cache and said file management table being necessary to write back to the physical disk device.

25 17. The file system as claimed in claim 14, wherein

a disk controller in said physical disk device comprises:

a disk cache for temporarily holding data to be transferred to and from a disk area; and

means for writing back data stored in a disk cache
5 provided in another disk controller in said physical disk device to the disk area;

wherein during a process of switching the I/O paths, a disk controller included in the new operational I/O path and provided in the physical disk device writes back data
10 selected from data stored in a disk cache in another disk controller included in said faulty I/O path and provided in said physical disk device, to said physical disk device, which was being accessed using said currently faulty I/O path, said selected data being necessary to write back to
15 said physical disk device.

18. The file system as claimed in claim 14, wherein at the time of completing a process of switching the I/O paths, the file server in the node to which the disk device storing the mount construction file is connected updates
20 said mount construction file, and rewrites availability information on the currently unavailable operational I/O path by information indicating "unavailable".

19. The file system as claimed in claim 5, wherein said mount construction file includes availability
25 information which is set for each I/O path and indicates

whether the I/O path is available, and said logical disk management table includes status flags each indicating an operational state of each I/O path registered therein,

whereby at system start-up, the file server which
5 performs said mount operation sets "operational" for each
of status flags in the logical disk management table
corresponding to I/O paths whose availability information
is set to "available" in said mount construction file and
sets "unavailable" for each of status flags in the logical
10 disk management table corresponding to I/O paths whose
availability information is set to "unavailable" in said
mount construction file, and the file server performs a
file mirroring operation on physical disk devices
accessible from I/O paths whose status flag is set to
15 "operational" in said logical disk management table in
normal operation.

20. The file system as claimed in claim 19, wherein
when a fault has occurred in one of said operational paths,
a file server in a node which has detected said fault
20 updates a logical disk management table in the node by
setting "unavailable" for a status flag of said faulty I/O
path, and then copies contents of said logical disk
management table to logical disk management tables in all
other nodes, and the file server in said first node to
25 which the disk device storing the mount construction file

is connected updates said mount construction file, and rewrites availability information on said faulty I/O path by information indicating "unavailable" to disconnect said faulty I/O path.

5 21. A file system including at least one node having a file server which processes files distributed and managed in a plurality of physical disk devices, said files each having a defined file ID, wherein upon receiving a request for accessing a file specifying a file ID, said file server
10 obtains a logical disk ID for accessing said file based on a file management table, obtains an I/O path corresponding to the logical disk ID based on a logical disk management table, and accesses a physical disk device by use of said I/O path, and whereas when a fault has occurred in an
15 operational I/O path, the file server switches from the faulty operational I/O path to another I/O path by rewriting logical disk management tables in all nodes.

 22. The file system claimed in claim 21, wherein said file server writes back data to said physical disk
20 device by use of said another I/O path to which the I/O-path switching has been made, said data being originally stored in said physical disk device.

 23. An I/O-path switching method employed in a file system which includes at least one node having a file
25 server which processes files distributed and managed in a

plurality of physical disk devices, said I/O-path switching method being performed by said server and comprising the steps of:

upon receiving a request for accessing a file
5 specifying a file ID from a user, obtaining a logical disk ID for accessing said file based on a file management table;

obtaining an I/O path corresponding the logical disk ID based on a logical disk management table;

10 accessing a physical disk device by use of said I/O path; and

when a fault has occurred in an operational I/O path, rewriting logical disk management tables in all nodes to switch the I/O path.

15